

# Dorling Kindersley

## BUSINESS CHANGE BRINGS CHALLENGES

The professional services team of Copyright Clearance Center (CCC) worked with the IT leadership team at DK to develop a custom content platform and search interface to support the wider strategic goals of DK. The project's mission was to enable powerful granular content discovery across the company's internal product store for more efficient content reuse, easier creation of new content products and improved support for additional revenue programs. The company's shift from print-led to digitally-driven content products solidified the need for a new, more robust system.

Having previously worked together, the CCC and DK team leveraged a proven phased approach to bring multiple product archives holding over 5,000 unique products – estimated at over 70 terabytes of data – together into a single, comprehensive archive. This archive would benefit the business through faster content discoverability, improved reuse, and better management of workflows for product archiving.

The original DK product store ran on an outdated version of Documentum document management software. The DK team was unhappy with the slow speed, non-intuitive interface and limited search capabilities in this particular implementation. The system was unable to display relationships between content products, so internal users had trouble finding material and frequently had to put in requests for manual assistance. Workflow processes were not documented or audited, leading to inconsistency in their application, unpredictable delivery of requests, a lack of transparency in project status and overall user frustration.

System limitations meant the DK team struggled to leverage its most valuable asset — the company's extensive library of unique content — leading to costly inefficiencies in product assembly, delays in time-to-market, and difficulty filling customer demand.

"At the time of this project, we were in transition of how we were thinking about content beyond print," said Russell Watts, Program Manager at DK. "CCC helped us in that transitional stage and beyond, offering some great expertise in terms of technology as well as change management and transformation projects."



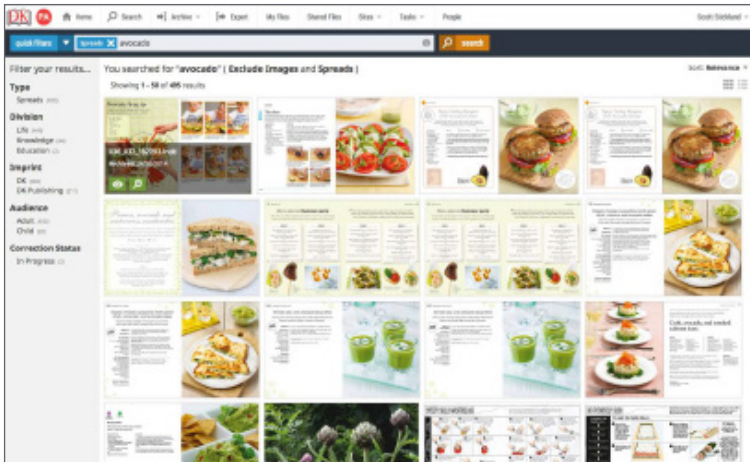
### About Dorling Kindersley

Established in 1974, Dorling Kindersley (DK) publishes best-selling, award-winning and visually stunning content for all audiences, in which the unrivalled clarity of the words and pictures come together to spectacular effect. DK Publishing is world renowned for its distinctive, highly visual books that inform, inspire, and entertain readers of all ages. Publisher of New York Times bestselling LEGO® and Star Wars® non-fiction books, Smithsonian titles, the award-winning Eyewitness series for children and Eyewitness Travel Guides, in addition to a wide selection of other books for adults and children. BradyGames, Alpha (Idiot's Guides) and Rough Guides are also available from DK, a division of Penguin Group.

After DK merged with Penguin Random House, the company experienced a period of significant change as control over various processes and technology shifted to DK. With the merger, plans already in place to address existing inefficiencies in the company's legacy content store were quickly accelerated to support the needs and address the potential of the combined merged entity.

### Technologies used

- Alfresco Web Framework
- ElasticSearch
- Amazon Web Services incl. EC2, RDS, S3, Cloudfront and SQS
- OWLIM RDF framework
- Mongo DB



## SOLUTION

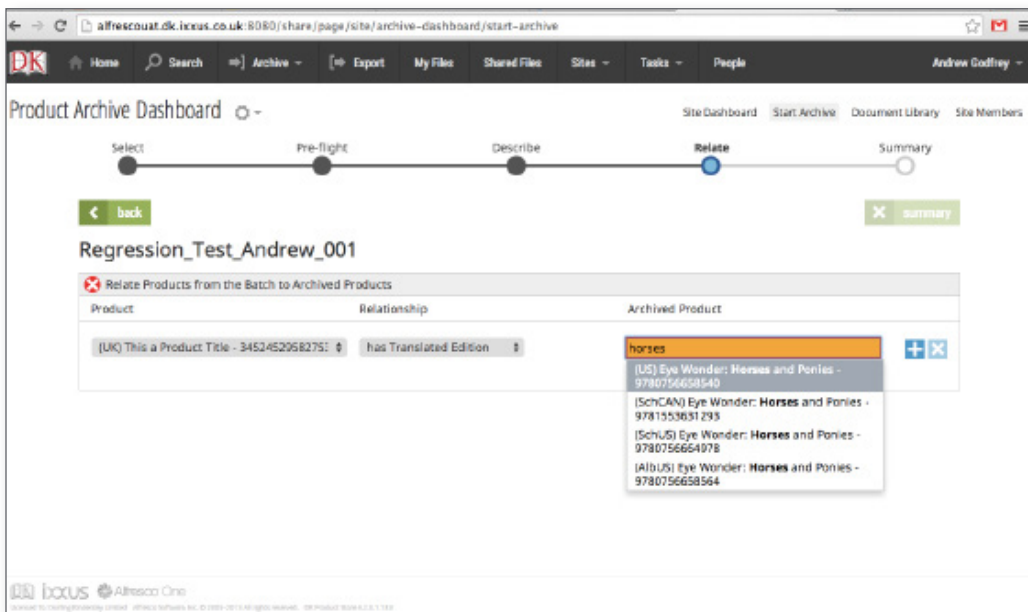
Created by the CCC professional services team in collaboration with DK, the new DK Product Archive is a cloud-based custom application used by internal teams to manage archived content. The Product Archive system allows production and creative teams to archive content, store updates and corrections with the source material and reuse content for new products, all through a consistent user interface specifically tailored to DK's unique requirements.

The Product Archive solution is hosted on Amazon Web Services (AWS) within a secure Virtual Private Cloud (VPC). It also uses:

- The Alfresco web framework and workflow engine for archive management, notifications, corrections, content requests and export
- Elasticsearch as the primary search service
- OWLIM as the semantic repository, combining RDF triple stores, inference engine and SPARQL query engine to form meaningful relationships between different products, their content and assets
- MongoDB as a source of audit and user activity information
- Amazon SQS service as the robust message queue to manage communication between the technology components, and
- Amazon S3 for scalable cloud storage of physical files such as IDML files.

The Product Archive solution also leverages various PRH services including Active Directory for user authentication, SMTP relay for email notifications and Local Network for file transfers.

The team created the ability to support the massive migration of content into the new solution, an approach that both reflected the existing content structure and also reflected the need to add information to drive content discoverability and user experience.



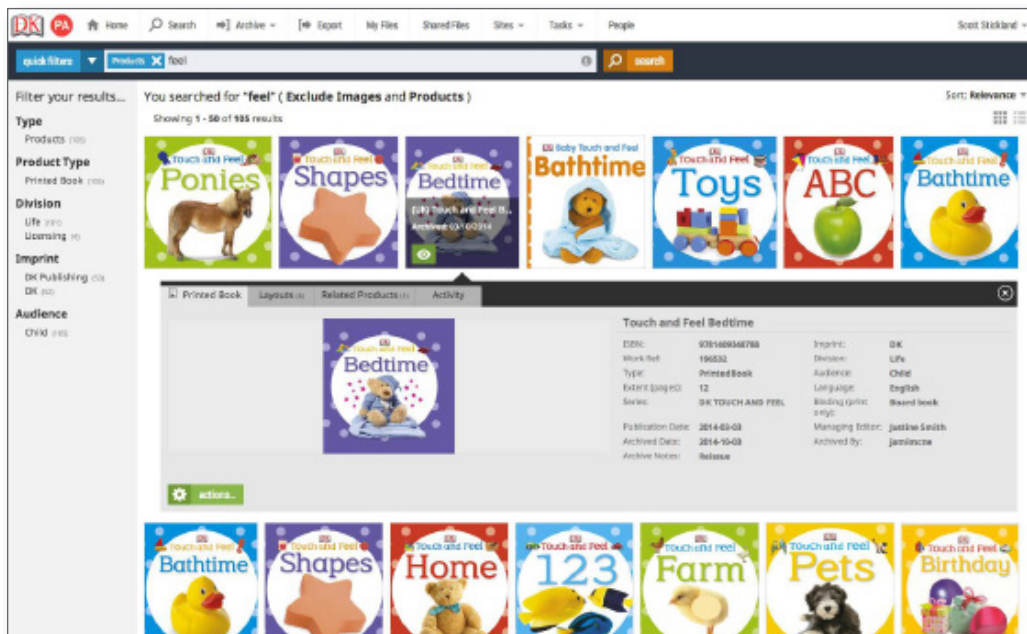
## Enhanced Discoverability

With metadata enrichment, semantic services and a big data store, the DK Product Archive solution delivered scalable discoverability and management of the company's print-oriented content assets while facilitating the transition to more digital content. Powerful search capabilities, including custom filters, facets and ISBN search, made it simple for DK users to surface the most relevant content to their needs and get the most value from each content asset.



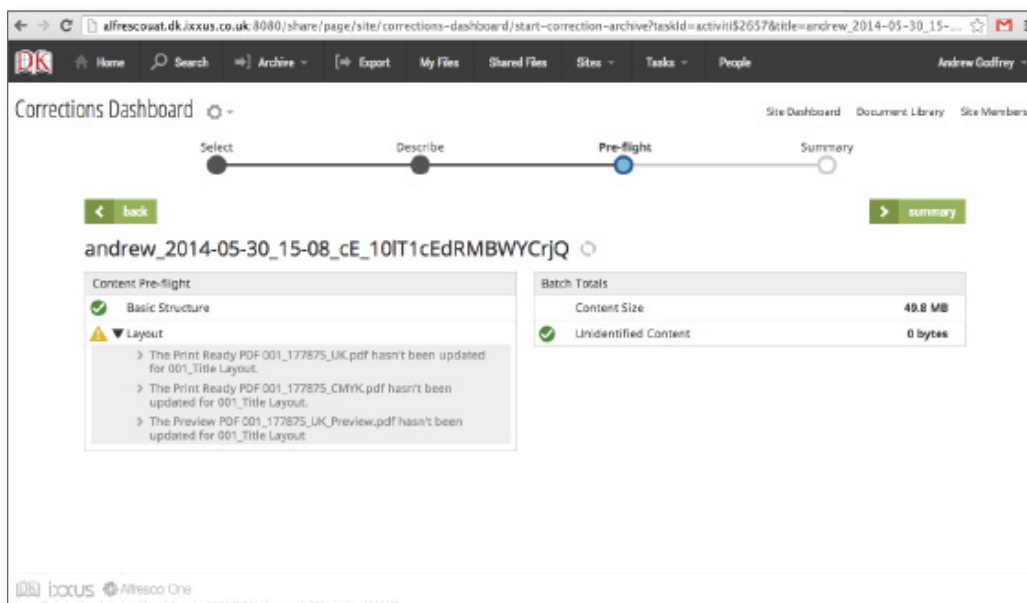
## Sleek & Visual User Interface

DK's products are known for their eye-catching design. As a visually led, design-oriented organization, DK requested at-a-glance access to images and full-page spreads. By incorporating thumbnail display, multi-page layouts and easily navigable product ontologies, the Product Archive solution allows a user to get high-level, image-driven views across DK's full product archive.



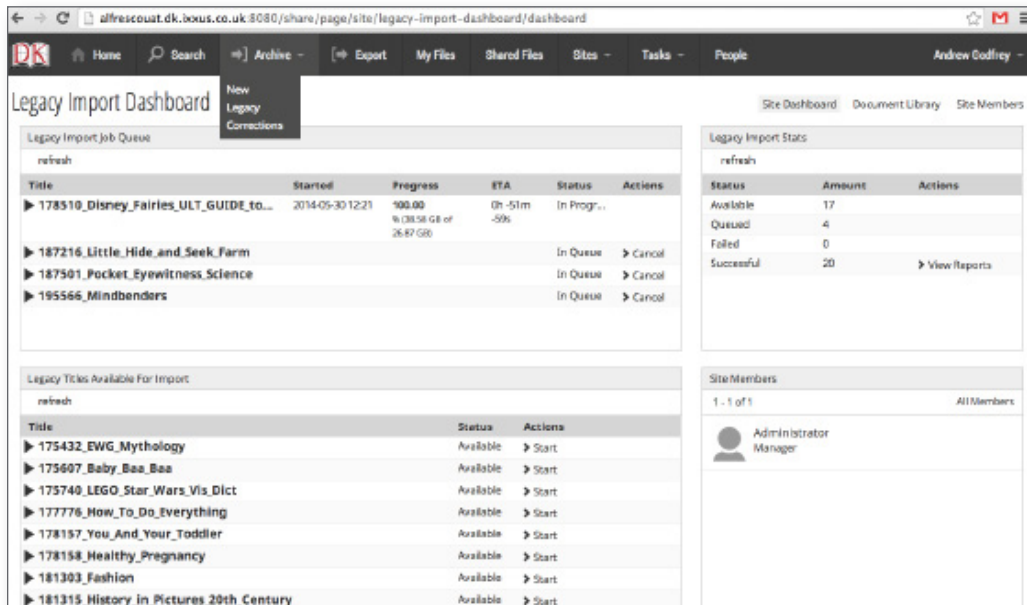
## Full Content Lifecycle

The Product Archive solution also allows control of the full renditions and corrections lifecycle, like highlighting whether a particular content asset has been overused, or is only licensed for particular types of use, or has been translated into linguistic variants (e.g. US/ UK editions). With an integrated corrections dashboard in the Product Archive, DK staff experience fewer data quality issues, making it easier to meet the high standards consumers have come to expect.



## Image Transformation for Thumbnail Preview

By integrating automated image transformation in the background, the Product Archive solution renders large files such as IDMLs, AIs, EPSs and PSDs as smaller, more accessible PDFs, JPGs and PNGs to allow for in-platform thumbnail view. Users can view high quality thumbnail images and multi-page spread previews as part of search results, allowing them to quickly make informed editorial and production decisions without having to first access large, multi-asset files. This accelerates product assembly and curation, driving editorial efficiencies and speeding time-to-market.



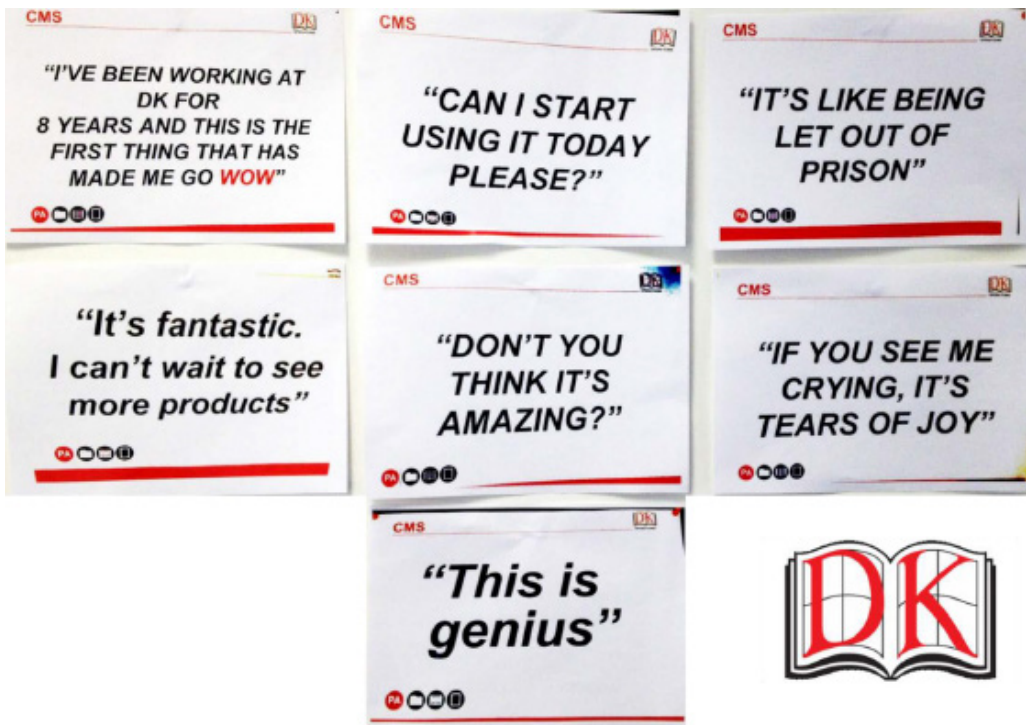
The screenshot displays the Legacy Import Dashboard interface. It features a navigation bar with options like Home, Search, Archive, Expert, My Files, Shared Files, Sites, Tasks, and People. The main content area is divided into several sections:

- Legacy Import Job Queue:** A table with columns for Title, Started, Progress, ETA, Status, and Actions. It lists three items: '178510\_Disney\_Fairies\_ULI\_GUIDE\_to...', '187216\_Little\_Hide\_and\_Seek\_Farm', and '187501\_Pocket\_Eyewitness\_Science'.
- Legacy Import Status:** A summary table with columns for Status, Amount, and Actions. It shows counts for Available (17), Queued (4), Failed (0), and Successful (20).
- Legacy Titles Available For Import:** A table with columns for Title, Status, and Actions. It lists titles such as '175432\_EWG\_Mythology', '175607\_Baby\_Bee\_Bee', and '175740\_LEGO\_Star\_Wars\_Vis\_Dict'.
- Site Members:** A section showing the current user as 'Administrator Manager'.

## RESULTS

Over a four-phase project, the CCC team successfully delivered the Product Archive solution as defined by DK's stakeholders. Each development phase including AWS infrastructure setup was delivered on time and on budget, consistently augmenting the original scope with key improvements in the true sense of Agile development. Key stakeholders are pleased with the intuitive user interface and functionality.

The Product Archive went live to a core set of DK users before being rolled out to the 700+ member creative teams across the globe. The response by stakeholders included these comments:



## BENEFITS OF THE PRODUCT ARCHIVE SOLUTION

### ① Content Discoverability and Reuse

Enhancing the discoverability and accessibility of content and assets allows archived products to find new relevancy and new leases on life. Once discovered, legacy content can also be repurposed, remixed and reused to create entirely new product, opening up potential revenue opportunities.

### ② Consistently Managed Workflow

All content processing in and out of the solution is wrapped with a consistent workflow definition that mirrors DK's processes, allowing a centralized overview for content-processing teams and enabling efficient information handovers. The improved, highly discoverable content store in Product Archive offers streamlined production, enabling major efficiency gains and an intuitive, user-driven experience.

### ③ Improved analytics capability

By building a foundation for future analytics and reporting, the Product Archive solution will deliver critical business intelligence within 1-2 years, allowing DK to discover things like: how many products they have on X subject, how many versions exist of X publication, how many times has X image been used cross-publication, etc.

### ④ Cost and Control

Cost and control were significant issues for DK, with a need to fulfill requests immediately after being made, even when the system was under heavy load. By deploying the entire Product Archive solution out to AWS managed services, DK now has direct control over scalability and development of the product store, allowing them to own their content like never before.

## THE FUTURE

After completing the initiative, DK and CCC entered into a managed service agreement for the Product Archive solution and developed plans to extend the current feature set, incorporate new content types and expand the solution's scope to include Work in Progress (WIP)/Authoring tools in order to achieve an end-to-end publishing and editorial solution. Under the managed services agreement, CCC provides application support, full systems management of the AWS environment and a proactive monitoring service.

### Digital transformation

DK and CCC also worked together to create an ongoing content roadmap and digital transformation strategy for the organization.

"With pressure to get content out of the old system and under our own control following the merger, time was of the essence," said Peter Buckley, Technology Director at DK. "Not only did CCC deliver on time, they went above and beyond the initial scope of the project. Our users and stakeholders have felt involved in the process since day one, building a close working relationship with the CCC team — something that shows in their eagerness to get their hands on the finished product!"



Copyright Clearance Center (CCC) builds unique solutions that connect content and rights in contextually relevant ways through software and professional services. CCC helps people navigate vast amounts of data to discover actionable insights, enabling them to innovate and make informed decisions. CCC, with its subsidiary RightsDirect, collaborates with customers to advance how data and information is integrated, accessed, and shared while setting the standard for effective copyright solutions that accelerate knowledge and power innovation. CCC is headquartered in Danvers, Mass. and has offices across North America, Europe and Asia. To learn more about CCC, visit [www.copyright.com](http://www.copyright.com).



### LEARN MORE

-  1.978.750.8400 (option 3)
-  [info@copyright.com](mailto:info@copyright.com)
-  [copyright.com](http://copyright.com)